

# On Score-Functions and Goodness-of-Fit Tests for Stochastic Processes

Yu.A. Kutoyants

Laboratoire de Statistique et Processus, Université du Maine  
Le Mans, France

and

Laboratory of Quantitative Finance, Higher School of Economics  
Moscow, Russia

## Abstract

The problems of the construction of the asymptotically distribution free goodness-of-fit tests for three models of stochastic processes are considered. The null hypothesis for all models is composite parametric. All tests are based on the score-function processes, where the unknown parameter is replaced by the MLE. We show that a special change of time transforms the limit score-function processes into the Brownian bridge. This property allows us to construct the asymptotically distribution free tests for the following three models of stochastic processes : dynamical systems with small noise, ergodic diffusion processes, inhomogeneous Poisson processes and nonlinear AR time series.

MSC 2000 Classification: 62M02, 62G10, 62G20.

*Key words:* Cramér-von Mises type tests, dynamical systems, small noise, ergodic diffusion process, inhomogeneous Poisson processes, nonlinear AR, goodness-of-fit tests, asymptotically distribution free tests.

## 1 Introduction

We consider the problem of the construction of asymptotically distribution free goodness-of-fit tests for the three models of stochastic processes observed in continuous time: small noise diffusion, ergodic diffusion and inhomogeneous Poisson process. We assume that under the basic hypotheses the models depend on some unknown one-dimensional parameter.

Let us recall what happens in the similar problem in the well-known i.i.d. model. Suppose that we observe  $n$  i.i.d. r.v.'s  $(X_1, \dots, X_n) = X^n$  with continuous distribution function  $F(x)$  and the basic (null) hypothesis is parametric

$$\mathcal{H}_0 : F(x) = F(\vartheta, x), \quad \vartheta \in \Theta$$

where  $F(\vartheta, x)$  is known smooth function of  $\vartheta \in \Theta = (a, b)$  and  $x$ .

We have to construct a goodness-of-fit (GoF) test  $\hat{\psi}_n$  which belongs to the class  $\mathcal{K}_\alpha$  of tests of asymptotic size  $\alpha$ , i.e.,

$$\mathcal{K}_\alpha = \{\bar{\psi}_n : \mathbf{E}_{\vartheta} \bar{\psi}_n = \alpha + o(1)\} \quad \text{for all } \vartheta \in \Theta.$$

Introduce the Cramér-von Mises type statistic

$$\delta_n = n \int_{-\infty}^{\infty} [\hat{F}_n(x) - F(\hat{\vartheta}_n, x)]^2 dF(\hat{\vartheta}_n, x), \quad \hat{F}_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}_{\{X_j \leq x\}},$$

where  $\hat{\vartheta}_n$  is the maximum likelihood estimator (MLE) and  $\hat{F}_n(x)$  is the empirical distribution function.

Note that if  $\Theta = \{\vartheta_0\}$  (simple basic hypothesis), then

$$\begin{aligned} \delta_n &= n \int_{-\infty}^{\infty} [\hat{F}_n(x) - F(\vartheta_0, x)]^2 dF(\vartheta_0, x) \\ &\Rightarrow \int_{-\infty}^{\infty} B(F(\vartheta_0, x))^2 dF(\vartheta_0, x) = \int_0^1 B(s)^2 ds \equiv \Delta, \end{aligned}$$

where  $s = F_0(\vartheta, x)$  and  $B(s), 0 \leq s \leq 1$  is a Brownian bridge. Therefore the test  $\hat{\psi}_n = \mathbb{I}_{\{\delta_n > c_\alpha\}}$  where  $c_\alpha$  is the solution of equation  $\mathbf{P}(\Delta > c_\alpha) = \alpha$  belongs to  $\mathcal{K}_\alpha$ . Moreover it is *asymptotically distribution free* (ADF), because the limit distribution of the statistic  $\delta_n$  does not depend on  $F(\vartheta_0, \cdot)$ .

Let us return to the parametric basic hypothesis and suppose that the model is sufficiently regular to satisfy the presented below expansion of the MLE:

$$\begin{aligned} u_n(x) &= \sqrt{n} (\hat{F}_n(x) - F(\hat{\vartheta}_n, x)) \\ &= \sqrt{n} (\hat{F}_n(x) - F(\vartheta, x)) + \sqrt{n} (F(\vartheta, x) - F(\hat{\vartheta}_n, x)) \\ &= B_n(x) - \sqrt{n} (\hat{\vartheta}_n - \vartheta) \dot{F}(\vartheta, x) + o(1). \end{aligned}$$

Here  $\dot{F}(\vartheta, x)$  means the derivative of  $F(\vartheta, x)$  w.r.t.  $\vartheta$ . The first term  $B_n(x) = \sqrt{n}(\hat{F}_n(x) - F(\vartheta, x))$  as before converges to the Brownian bridge  $B(F(\vartheta, x))$  and the MLE admits the representation

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{\dot{\ell}(\vartheta, X_j)}{I(\vartheta)} + o(1) = \int \frac{\dot{\ell}(\vartheta, y)}{I(\vartheta)} dB_n(y) + o(1).$$

Here  $\ell(\vartheta, x) = \ln f(\vartheta, x)$ ,  $f(\vartheta, x)$  is the density function and  $I(\vartheta)$  is the Fisher information. It can be shown that

$$\begin{aligned} u_n(x) &\Rightarrow B(F(\vartheta, x)) - \int \frac{\dot{\ell}(\vartheta, y)}{\sqrt{I(\vartheta)}} dB(F(\vartheta, y)) \int_{-\infty}^x \frac{\dot{\ell}(\vartheta, y)}{\sqrt{I(\vartheta)}} dF(\vartheta, y) \\ &= B(s) - \int_0^1 h(\vartheta, v) dB(v) \int_0^s h(\vartheta, v) dv \equiv u(s), \end{aligned} \quad (1)$$

where  $s = F(\vartheta, x)$ ,

$$h(\vartheta, s) = \frac{\dot{\ell}(\vartheta, F_{\vartheta}^{-1}(s))}{\sqrt{I(\vartheta)}}, \quad \int_0^1 h(\vartheta, v)^2 dv = 1.$$

Therefore  $u_n(\cdot)$  converges to the random function  $u(\cdot)$  and this allows us to prove (see Darling [2]) the convergence

$$\delta_n \Rightarrow \int_0^1 u(s)^2 ds.$$

Hence the test based on  $\delta_n$  is not ADF because the limit distribution of the statistic  $\delta_n$  depends on  $F(\vartheta, x)$ . This makes the choice of the threshold  $c_\alpha$  a more difficult problem.

One possibility to obtain ADF test is to find a linear transformation of  $u(\cdot)$  into Wiener process:  $L_W[u](s) = w(s)$ . Then

$$\int_{-\infty}^{\infty} \left( L_W[u](F(\vartheta, x)) \right)^2 dF(\vartheta, x) = \int_0^1 w(s)^2 ds \equiv \hat{\delta}.$$

Therefore if we take the statistics

$$\hat{\delta}_n = \int_{-\infty}^{\infty} \left( L_W[u_n](F(\hat{\vartheta}_n, x)) \right)^2 dF(\hat{\vartheta}_n, x)$$

and verify the convergence  $\hat{\delta}_n \Rightarrow \hat{\delta}$ , then the test  $\hat{\psi}_n = \mathbb{I}_{\{\hat{\delta}_n > d_\alpha\}}$  with  $\mathbf{P}(\hat{\delta} > d_\alpha) = \alpha$  is ADF and belongs to  $\mathcal{K}_\alpha$ . Note that such transformation  $L_W[u]$  was proposed by Khmaladze [8] (see also the different proof of it in [9]).

In the present work we consider a similar problem of construction of ADF GoF tests for stochastic processes, for which we suggest a much simpler transformation of the corresponding limit statistics into the Brownian bridge.

The goal of this work is to study the GoF tests for three models of observations of continuous time stochastic processes: diffusion processes  $X^\varepsilon = (X_t, 0 \leq t \leq T)$  with small diffusion coefficient ( $\varepsilon \rightarrow 0$ ), ergodic diffusion processes  $X^T = (X_t, 0 \leq t \leq T)$ ,  $T \rightarrow \infty$  and  $\tau_*$ -periodic Poisson processes  $X^n = (X_t, 0 \leq t \leq T = \tau_* n)$ ,  $n \rightarrow \infty$ . For all three models we introduce the corresponding score-function processes (SFP)  $U_\varepsilon(\cdot)$ ,  $U_T(\cdot)$  and  $U_n(\cdot)$  and then we show that the Cramér-von Mises type statistics based on these SFP allow us to construct the ADF GoF tests as follows. We also discuss the possibility of construction of similar tests in the case of i.i.d. observations and in the case of nonlinear AR time series.

First we show that the corresponding SFP's  $U_\varepsilon(\cdot)$ ,  $U_T(\cdot)$  and  $U_n(\cdot)$  converge to the processes  $(U(\vartheta, t), 0 \leq t \leq T)$ ,  $(U(\vartheta, x), x \in R)$  and  $(U(\vartheta, t), 0 \leq t \leq \tau_*)$  respectively. Say,  $U_\varepsilon(\cdot)$  converges to

$$U(\vartheta, t) = \int_0^t h(s) dW_s - \int_0^T h(s) dW_s \int_0^t h(s)^2 ds, \quad \int_0^T h(s)^2 ds = 1,$$

where  $h(s) = h(\vartheta, s)$  is some function and  $W_s, 0 \leq s \leq T$  is a Wiener process. Therefore if we put

$$\tau = \int_0^t h(\vartheta, s)^2 ds, \quad \int_0^t h(s) dW_s = W\left(\int_0^t h(\vartheta, s)^2 ds\right) = W(\tau),$$

where  $W(\cdot)$  is another Wiener process, then we can write

$$U(\vartheta, t) = W(\tau) - W(1) \tau = B(\tau), \quad 0 \leq \tau \leq 1,$$

where  $B(\cdot)$  is a Brownian bridge. Hence

$$\int_0^T U(\vartheta, t)^2 h(\vartheta, t)^2 dt = \int_0^1 B(\tau)^2 d\tau = \Delta.$$

This suggests the construction of tests with the help of “empirical versions”  $U_{\varepsilon, T, n}(\cdot)$  and  $h_{\varepsilon, T, n}(\cdot)$  of  $U(\cdot)$  and  $h(\cdot)$  as follows. Introduce the corresponding statistics (symbolic writing)

$$\Delta_{\varepsilon, T, n} = \int U_{\varepsilon, T, n}(s)^2 h_{\varepsilon, T, n}(s)^2 ds.$$

Then we show that for all three models we have the convergences to the same limit

$$\Delta_\varepsilon \Longrightarrow \Delta, \quad \Delta_T \Longrightarrow \Delta, \quad \Delta_n \Longrightarrow \Delta$$

and therefore the tests

$$\hat{\psi}_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}, \quad \hat{\psi}_T = \mathbb{I}_{\{\Delta_T > c_\alpha\}}, \quad \hat{\psi}_n = \mathbb{I}_{\{\Delta_n > c_\alpha\}}, \quad \mathbf{P}(\Delta > c_\alpha) = \alpha$$

are ADF. Below we realize this program. Moreover we show that this approach cannot be applied directly to the model of observations of i.i.d. random variables, but in the case of nonlinear AR time series we have the similar ADF GoF test, of course, under the strong regularity conditions.

This work is a continuation of the study of GoF tests for diffusion processes observed in continuous time. The case of simple basic hypothesis was treated for example in the works [4],[7], [12], [1], [20], [14]. The case of parametric basic hypothesis and ADF tests was studied in the works [21], [14], [9], [15], [16].

For point processes there are many publications devoted to this subject, see, e.g., [19] and the references therein.

## 2 Score-Function Processes

We have three stochastic processes observed in continuous time : small noise diffusion, ergodic diffusion and inhomogeneous Poisson processes. First we consider limits of the SFP's, separately for these models of observations. Then we show how these limits can be used for construction of the ADF GoF tests.

### 2.1 Small Noise Diffusion Processes.

We observe a realization  $X^\varepsilon = (X_t, 0 \leq t \leq T)$  of diffusion process satisfying the stochastic differential equation

$$dX_t = S(t, X_t) dt + \varepsilon \sigma(t, X_t) dW_t, \quad x_0, \quad 0 \leq t \leq T, \quad (2)$$

where the trend coefficient  $S(t, X_t)$  is an unknown function and the diffusion coefficient  $\varepsilon^2 \sigma(t, X_t)^2$  is a known positive function. The initial value  $x_0$  is deterministic and  $\varepsilon \in (0, 1]$ .

We have to test the following parametric (basic) hypothesis:

$\mathcal{H}_0$  : *The observed process has the stochastic differential*

$$dX_t = S(\vartheta, t, X_t) dt + \varepsilon \sigma(t, X_t) dW_t, \quad x_0, \quad 0 \leq t \leq T, \quad (3)$$

where the trend coefficient  $S(\vartheta, t, X_t)$  is a known smooth function depending on some unknown parameter  $\vartheta \in \Theta = (a, b)$ .

Our goal is to construct a GoF test  $\hat{\psi}_\varepsilon$ , which belongs to the class  $\mathcal{K}_\alpha$  and is consistent in the asymptotics of *small noise*  $\varepsilon \rightarrow 0$ . Note that this stochastic model and the statistical inference for it has been considered in many works. See, for example, [5], [10] [22] and the references therein.

Let us introduce the following regularity condition.

$\mathcal{R}$ . *The functions  $S(\vartheta, t, x)$  and  $\sigma(t, x)$  have two continuous bounded derivatives with respect to  $\vartheta$  and  $x$  and have continuous bounded derivatives w.r.t.  $t$ .*

Below the dot stands for the derivative w.r.t.  $\vartheta$  and prime means the derivative w.r.t.  $x$  or w.r.t.  $t$ . For example,

$$\ddot{S}(\vartheta, t, x) = \frac{\partial^2 S(\vartheta, t, x)}{\partial \vartheta^2}, \quad S'_x(\vartheta, t, x) = \frac{\partial S(\vartheta, t, x)}{\partial x}.$$

Let us denote by  $x^T = (x_t, 0 \leq t \leq T)$  the solution of the equation (3) with  $\varepsilon = 0$ , i.e.  $x^T$  is solution of the ordinary differential equation

$$\frac{dx_t}{dt} = S(\vartheta, t, x_t), \quad x_0, \quad 0 \leq t \leq T.$$

Of course it is a function of  $\vartheta$ , i.e.  $x_t = x_t(\vartheta)$ . It is known that as  $\varepsilon \rightarrow 0$ , the process  $X^\varepsilon$  converges to the deterministic function  $x^T$  and this convergence is uniform w.r.t.  $t \in [0, T]$  (see [5]).

Further, assume that the following identifiability condition is fulfilled.

$\mathcal{I}$ . *For any  $\nu > 0$*

$$\inf_{\vartheta_0 \in \Theta} \inf_{|\vartheta - \vartheta_0| > \nu} \int_0^T \left[ \frac{S(\vartheta, t, x_t^*) - S(\vartheta_0, t, x_t^*)}{\sigma(t, x_t^*)} \right]^2 dt > 0.$$

Here and below  $x_t^* = x_t(\vartheta_0)$ .

The likelihood ratio function in the case of observations (3) is

$$L(\vartheta, X^\varepsilon) = \exp \left\{ \int_0^T \frac{S(\vartheta, t, X_t)}{\varepsilon^2 \sigma(t, X_t)^2} dX_t - \int_0^T \frac{S(\vartheta, t, X_t)^2}{2\varepsilon^2 \sigma(t, X_t)^2} dt \right\}, \quad \vartheta \in \Theta$$

and the MLE  $\hat{\vartheta}_\varepsilon$  is defined by the equation

$$L(\hat{\vartheta}_\varepsilon, X^\varepsilon) = \sup_{\vartheta \in \Theta} L(\vartheta, X^\varepsilon). \quad (4)$$

The MLE  $\hat{\vartheta}_\varepsilon$  under the aforementioned regularity conditions admits the representation

$$\varepsilon^{-1} (\hat{\vartheta}_\varepsilon - \vartheta) = \mathbf{I}(\vartheta)^{-1} \int_0^T \frac{\dot{S}(\vartheta, t, x_t)}{\sigma(t, x_t)} dW_t + o(1) \quad (5)$$

see [10]. Here  $I(\vartheta)$  is the Fisher information

$$I(\vartheta) = \int_0^T \left( \frac{\dot{S}(\vartheta, t, x_t)}{\sigma(t, x_t)} \right)^2 dt > 0.$$

We define the score-function

$$\frac{\partial \ln L(\vartheta, X^\varepsilon)}{\partial \vartheta} = \int_0^T \frac{\dot{S}(\vartheta, t, X_t)}{\varepsilon^2 \sigma(t, X_t)^2} [dX_t - S(\vartheta, t, X_t) dt]$$

and the normalized score-function

$$U_\varepsilon(\vartheta, X^\varepsilon) = \int_0^T \frac{\dot{S}(\vartheta, t, X_t)}{\varepsilon I(\vartheta)^{1/2} \sigma(t, X_t)^2} [dX_t - S(\vartheta, t, X_t) dt].$$

If the true value is  $\vartheta_0$ , then we have the convergence

$$U_\varepsilon(\vartheta_0, X^\varepsilon) = \int_0^T \frac{\dot{S}(\vartheta_0, t, X_t)}{I(\vartheta_0)^{1/2} \sigma(t, X_t)} dW_t \longrightarrow \zeta,$$

where

$$\zeta = \int_0^T \frac{\dot{S}(\vartheta_0, t, x_t^*)}{I(\vartheta_0)^{1/2} \sigma(t, x_t^*)} dW_t \sim \mathcal{N}(0, 1).$$

The proof, which can be found in [10], follows from the uniform convergence of  $X_t$  to  $x_t^*$ .

Let us introduce the score-function process

$$U_\varepsilon(t, \vartheta, X^\varepsilon) = I(\vartheta)^{-1/2} \int_0^t \frac{\dot{S}(\vartheta, s, X_s)}{\varepsilon \sigma(s, X_s)^2} [dX_s - S(\vartheta, s, X_s) ds], \quad 0 \leq t \leq T,$$

and (formally) the statistic  $U_\varepsilon(t) = U_\varepsilon(t, \hat{\vartheta}_\varepsilon, X^\varepsilon)$ ,  $0 \leq t \leq T$ . We say “formally” because the MLE  $\hat{\vartheta}_\varepsilon$  depends on the whole trajectory  $X^\varepsilon$  and the corresponding Itô integral

$$\int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\sigma(s, X_s)^2} dX_s \tag{6}$$

is not well defined. The correct definition will be given later and here we show (as well formally) to which limit this process can be expected to converge.

Note that  $U_\varepsilon(T, \vartheta_0, X^\varepsilon) = U_\varepsilon(\vartheta_0, X^\varepsilon)$  with  $\mathbf{P}_{\vartheta_0}^{(\varepsilon)}$  probability 1.

We have ( $\vartheta_0$  is the true value)

$$\begin{aligned}
U_\varepsilon(t) &= \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon \mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} \left[ dX_s - S(\hat{\vartheta}_\varepsilon, s, X_s) ds \right] \\
&= \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)} dW_s \\
&\quad + \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) \left[ S(\vartheta_0, s, X_s) - S(\hat{\vartheta}_\varepsilon, s, X_s) \right]}{\varepsilon \mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} ds \\
&= \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)} dW_s \\
&\quad - \frac{\hat{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon \mathbb{I}(\hat{\vartheta}_\varepsilon)^{-1/2}} \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) \dot{S}(\tilde{\vartheta}_\varepsilon, s, X_s)}{\mathbb{I}(\hat{\vartheta}_\varepsilon) \sigma(s, X_s)^2} ds \\
&= \int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)}{\mathbb{I}(\vartheta_0)^{1/2} \sigma(s, x_s^*)^2} dW_s \\
&\quad - \int_0^T \frac{\dot{S}(\vartheta_0, s, x_s^*)}{\mathbb{I}(\vartheta_0)^{1/2} \sigma(s, x_s^*)^2} dW_s \int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)^2}{\mathbb{I}(\vartheta_0) \sigma(s, x_s^*)^2} ds + o(1). \quad (7)
\end{aligned}$$

Further, if we denote

$$\tau = \int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)^2}{\mathbb{I}(\vartheta_0) \sigma(s, x_s^*)^2} ds, \quad 0 \leq \tau \leq 1,$$

then we can write

$$\int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)}{\mathbb{I}(\vartheta_0)^{1/2} \sigma(s, x_s^*)^2} dW_s = W(\tau),$$

where  $W(\cdot)$  is some Wiener process. Therefore we obtain the limit

$$U_\varepsilon(t) \implies W(\tau) - W(1)\tau = B(\tau), \quad 0 \leq \tau \leq 1,$$

with a Brownian bridge  $B(\cdot)$ .

This convergence suggests the construction of the following test statistic

$$\Delta_\varepsilon = \int_0^T \frac{U_\varepsilon(t)^2 \dot{S}(\hat{\vartheta}_\varepsilon, t, X_t)^2}{\mathbb{I}(\hat{\vartheta}_\varepsilon) \sigma(t, X_t)^2} dt \quad (8)$$

and the test  $\hat{\psi}_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}$ , where  $\mathbf{P}(\Delta > c_\alpha) = \alpha$ . If we verify that

$$\Delta_\varepsilon \implies \Delta = \int_0^1 B(\tau)^2 d\tau,$$



then the test  $\hat{\psi}_\varepsilon \in \mathcal{K}_\alpha$  and is ADF.

To avoid the problem concerning the stochastic integral (6) we use two possibilities: one is the well-known device which consists in the application of the Itô formula to the function

$$H(\vartheta, s, x) = \int_{x_0}^x \frac{\dot{S}(\vartheta, s, y)}{\sigma(s, y)^2} dy$$

and the second is based on some preliminary estimator of the parameter  $\vartheta$ .

The first approach was applied in the similar problem in [15] and here we follow the same steps. The second approach was mentioned in [15] too but here (below) we work out the details of the proof.

*The first approach.* The Itô formula applied to the function  $H(\vartheta, s, X_s)$  gives us the stochastic differential

$$\begin{aligned} \int_0^t \frac{\dot{S}(\vartheta, s, X_s)}{\sigma(s, X_s)^2} dX_s &= H(\vartheta, t, X_t) \\ &\quad - \int_0^t \left[ H'_s(\vartheta, s, X_s) + \frac{\varepsilon^2 \sigma(s, X_s)^2}{2} H''_{x,x}(\vartheta, s, X_s) \right] ds. \end{aligned}$$

Note that the contribution of the term

$$\varepsilon^2 \int_0^t \sigma(s, X_s)^2 H''_{x,x}(\vartheta, s, X_s) ds$$

is asymptotically negligible and we can omit it.

We have

$$\begin{aligned} \hat{U}_\varepsilon(t) &= \frac{H(\hat{\vartheta}_\varepsilon, t, X_t)}{\varepsilon I(\hat{\vartheta}_\varepsilon)^{1/2}} - \int_0^t \frac{H'_s(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon I(\hat{\vartheta}_\varepsilon)^{1/2}} ds - \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) S(\vartheta_0, s, X_s)}{\varepsilon I(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} ds \\ &\quad - \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) [S(\hat{\vartheta}_\varepsilon, s, X_s) - S(\vartheta_0, s, X_s)]}{\varepsilon I(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} ds + O(\varepsilon) \\ &= J(\hat{\vartheta}_\varepsilon, t, X^t) - K(\hat{\vartheta}_\varepsilon, t, X^t) + O(\varepsilon), \end{aligned} \tag{9}$$

where  $K(\cdot)$  is the last integral. Its convergence is obtained directly (see (5)):

$$\begin{aligned} K(\hat{\vartheta}_\varepsilon, t, X^t) &= \frac{\hat{\vartheta}_\varepsilon - \vartheta_0}{\varepsilon} \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) \dot{S}(\tilde{\vartheta}_\varepsilon, s, X_s)}{I(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} ds \\ &\longrightarrow \int_0^T \frac{\dot{S}(\vartheta_0, s, x_s^*)}{I(\vartheta_0)^{1/2} \sigma(s, x_s^*)} dW_s \int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)^2}{I(\vartheta_0) \sigma(s, x_s^*)^2} ds. \end{aligned}$$

Further, we verify that

$$J(\hat{\vartheta}_\varepsilon, t, X^t) - J(\vartheta_0, t, X^t) \longrightarrow 0$$

and that

$$J(\vartheta_0, t, X^t) \longrightarrow \int_0^t \frac{\dot{S}(\vartheta_0, s, x_s^*)}{\mathbb{I}(\vartheta_0)^{1/2} \sigma(s, x_s^*)} dW_s$$

(see details in [15]).

Thus we obtained the convergence mentioned in (7) and the following result.

**Proposition 1** *Suppose that the conditions of regularity are fulfilled, then the test  $\hat{\psi}_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}$  with*

$$\Delta_\varepsilon = \int_0^T \frac{\hat{U}_\varepsilon(t)^2 \dot{S}(\hat{\vartheta}_\varepsilon, t, X_t)^2}{\mathbb{I}(\hat{\vartheta}_\varepsilon) \sigma(t, X_t)^2} dt$$

*is ADF and belongs to  $\mathcal{K}_\alpha$ .*

*Second approach.* Let us write  $\hat{U}_\varepsilon(t)$  as the difference of two integrals

$$\hat{U}_\varepsilon(t) = \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon \mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} dX_s - \int_0^t \frac{\dot{S}(\hat{\vartheta}_\varepsilon, s, X_s) S(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon \mathbb{I}(\hat{\vartheta}_\varepsilon)^{1/2} \sigma(s, X_s)^2} ds.$$

Note that the properties of the estimator  $\hat{\vartheta}_\varepsilon$  required in the study of the first and the second integrals are different.

In the first integral it is sufficient that  $\hat{\vartheta}_\varepsilon \rightarrow \vartheta_0$  and in the second integral we need the asymptotic efficiency (full limit variance) of the MLE. Therefore we can consider two different estimators in the calculation of these integrals. For the first integral we introduce a preliminary (consistent) estimator  $\bar{\vartheta}_{\nu_\varepsilon}$  constructed by the first  $(X_t, 0 \leq t \leq \nu_\varepsilon)$  observations. Here  $\nu_\varepsilon \rightarrow 0$  but slowly. Then we can use the estimator  $\bar{\vartheta}_{\nu_\varepsilon}$  in the calculation of the integral

$$\int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s)}{\varepsilon \mathbb{I}(\bar{\vartheta}_{\nu_\varepsilon})^{1/2} \sigma(s, X_s)^2} dX_s, \quad t \in [\nu_\varepsilon, T],$$

which is now well defined. In the second integral we keep  $\hat{\vartheta}_\varepsilon$  in the function  $S(\hat{\vartheta}_\varepsilon, s, X_s)$  only. Therefore we consider the statistic

$$V_\varepsilon(t) = \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s)}{\varepsilon \mathbb{I}(\bar{\vartheta}_{\nu_\varepsilon})^{1/2} \sigma(s, X_s)^2} dX_s - \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s) S(\hat{\vartheta}_\varepsilon, s, X_s)}{\varepsilon \mathbb{I}(\bar{\vartheta}_{\nu_\varepsilon})^{1/2} \sigma(s, X_s)^2} ds,$$

where  $t \in [\nu_\varepsilon, T]$ . Now we can repeat the calculations similar to (7) for the statistic  $V_\varepsilon(t), t \in [\nu_\varepsilon, T]$ , which is this time well defined, and obtain the same limit expression.

Let us construct a consistent estimator  $\bar{\vartheta}_{\nu_\varepsilon}$  by the “vanishing observations”  $X_t, 0 \leq t \leq \nu_\varepsilon, \nu_\varepsilon \rightarrow 0$ . Introduce a minimum distance estimator (MDE)

$$\bar{\vartheta}_{\nu_\varepsilon} = \arg \inf_{\vartheta \in \Theta} \int_0^{\nu_\varepsilon} [X_t - x_t(\vartheta)]^2 dt.$$

The consistency of this estimator is verified in the following lemma.

**Lemma 1** *Suppose that the regularity condition  $\mathcal{R}$  is fulfilled and for all  $\vartheta \in \Theta$  we have  $\left| \dot{S}(\vartheta, 0, x_0) \right| \geq \kappa$ , where  $\kappa > 0$ . Then the MDE  $\bar{\vartheta}_{\nu_\varepsilon}$  with  $\nu_\varepsilon = \varepsilon^2 \ln(\varepsilon^{-1})$  is consistent.*

**Proof.** Below  $\|\cdot\|_{\nu_\varepsilon}$  is  $L^2[0, \nu_\varepsilon]$  norm. Let us put

$$g(\gamma, \nu_\varepsilon) = \inf_{|\vartheta - \vartheta_0| > \gamma} \|x_t(\vartheta) - x_t(\vartheta_0)\|_{\nu_\varepsilon}.$$

Note that

$$g(\gamma, \nu_\varepsilon)^2 = \int_0^{\nu_\varepsilon} [x_t(\vartheta) - x_t(\vartheta_0)]^2 dt = (\vartheta - \vartheta_0)^2 \int_0^{\nu_\varepsilon} \dot{x}_t(\tilde{\vartheta})^2 dt.$$

with some  $\tilde{\vartheta}$ . The derivative w.r.t.  $\vartheta$  of  $x_t(\vartheta)$  satisfies the equation

$$\frac{d\dot{x}_t(\vartheta)}{dt} = \dot{S}(\vartheta, t, x_t(\vartheta)) + S'_x(\vartheta, t, x_t(\vartheta)) \dot{x}_t(\vartheta), \quad \dot{x}_0(\vartheta) = 0.$$

Its solution is the function

$$\dot{x}_t(\vartheta) = \int_0^t \dot{S}(\vartheta, s, x_s(\vartheta)) \exp \left\{ \int_s^t S'_x(\vartheta, v, x_v(\vartheta)) dv \right\} ds.$$

Hence for the small values of  $t$  we have the estimate

$$\dot{x}_t(\vartheta) = t \dot{S}(\vartheta, 0, x_0) (1 + O(t)).$$

Therefore for all  $\varepsilon < \varepsilon_*$ , where  $\varepsilon_*$  is some small value

$$\|x_t(\vartheta) - x_t(\vartheta_0)\|_{\nu_\varepsilon}^2 \geq \frac{(\vartheta - \vartheta_0)^2 \kappa^2 \nu_\varepsilon^3}{6}.$$

Further, for any  $\gamma > 0$  we have

$$\begin{aligned}
& \mathbf{P}_{\vartheta_0} (|\bar{\vartheta}_{\nu_\varepsilon} - \vartheta_0| > \gamma) \\
&= \mathbf{P}_{\vartheta_0} \left( \inf_{|\vartheta - \vartheta_0| \leq \gamma} \|X_t - x_t(\vartheta)\|_{\nu_\varepsilon} > \inf_{|\vartheta - \vartheta_0| > \gamma} \|X_t - x_t(\vartheta)\|_{\nu_\varepsilon} \right) \\
&\leq \mathbf{P}_{\vartheta_0} \left( \inf_{|\vartheta - \vartheta_0| \leq \gamma} (\|X_t - x_t(\vartheta_0)\|_{\nu_\varepsilon} + \|x_t(\vartheta) - x_t(\vartheta_0)\|_{\nu_\varepsilon}) \right. \\
&\quad \left. > \inf_{|\vartheta - \vartheta_0| > \gamma} (\|x_t(\vartheta) - x_t(\vartheta_0)\|_{\nu_\varepsilon} - \|X_t - x_t(\vartheta_0)\|_{\nu_\varepsilon}) \right) \\
&= \mathbf{P}_{\vartheta_0} (2 \|X_t - x_t(\vartheta_0)\|_{\nu_\varepsilon} \geq g(\gamma, \nu_\varepsilon)) \\
&\leq \frac{4}{g(\gamma, \nu_\varepsilon)^2} \mathbf{E}_{\vartheta_0} \int_0^{\nu_\varepsilon} [X_t - x_t(\vartheta_0)]^2 dt \leq \frac{C\varepsilon^2\nu_\varepsilon^2}{\gamma^2\kappa^2\nu_\varepsilon^3} \leq \frac{C}{\ln \frac{1}{\varepsilon}} \rightarrow 0.
\end{aligned}$$

Here we used the estimate

$$\sup_{0 \leq s \leq t} \mathbf{E}_{\vartheta_0} |X_s - x_s(\vartheta_0)|^2 \leq Ct\varepsilon^2,$$

which can be found, for example, in [10], Lemma 1.13.

Therefore the estimator  $\bar{\vartheta}_{\nu_\varepsilon}$  is consistent and we have the following result.

**Proposition 2** *Suppose that the conditions of regularity are fulfilled and for all  $\vartheta \in \Theta$  we have  $|\dot{S}(\vartheta, 0, x_0)| \geq \kappa$ , where  $\kappa > 0$ , then the test  $\tilde{\psi}_\varepsilon = \mathbb{I}_{\{\tilde{\Delta}_\varepsilon > c_\alpha\}}$  with*

$$\tilde{\Delta}_\varepsilon = \int_{\nu_\varepsilon}^T \frac{V_\varepsilon(t)^2 \dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, t, X_t)^2}{I(\bar{\vartheta}_{\nu_\varepsilon}) \sigma(t, X_t)^2} dt$$

*is ADF and belongs to  $\mathcal{K}_\alpha$ .*

Let us consider the problem of *consistency* of this test. The observed process under alternative is

$$dX_t = S(t, X_t) dt + \varepsilon \sigma(t, X_t) dW_t, \quad X_0 = x_0, \quad 0 \leq t \leq T,$$

where  $S(t, x)$  does not belong to the parametric family of trend coefficients  $\{S(\vartheta, t, x), \vartheta \in \Theta\}$ . We obtain the following representation for the statistic

$V_\varepsilon(\cdot)$ :

$$\begin{aligned}
V_\varepsilon(t) &= \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s)}{I(\bar{\vartheta}_{\nu_\varepsilon})^{1/2} \sigma(s, X_s)^2} dW_s \\
&\quad + \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}_{\nu_\varepsilon}, s, X_s) [S(s, X_s) - S(\hat{\vartheta}_\varepsilon, s, X_s)]}{\varepsilon I(\bar{\vartheta}_{\nu_\varepsilon})^{1/2} \sigma(s, X_s)^2} ds \\
&= \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}, s, x_s)}{I(\bar{\vartheta})^{1/2} \sigma(s, x_s)^2} dW_s + o(1) \\
&\quad + \int_{\nu_\varepsilon}^t \frac{\dot{S}(\bar{\vartheta}, s, x_s) [S(s, x_s) - S(\hat{\vartheta}, s, x_s)]}{\varepsilon I(\bar{\vartheta})^{1/2} \sigma(s, x_s)^2} ds (1 + o(1)).
\end{aligned}$$

Here  $x_t$  is solution of the ordinary differential equation

$$\frac{dx_t}{dt} = S(t, x_t), \quad x_0, \quad 0 \leq t \leq T$$

and  $\hat{\vartheta}, \bar{\vartheta}$  are defined as follows

$$\hat{\vartheta} = \arg \inf_{\vartheta \in \Theta} \int_0^T \left( \frac{S(\vartheta, t, x_t) - S(t, x_t)}{\sigma(t, x_t)} \right)^2 dt, \quad (10)$$

$$\bar{\vartheta} = \arg \inf_{\vartheta \in \Theta} |S(\vartheta, 0, x_0) - S(0, x_0)|. \quad (11)$$

For the proof of (10) see [10], Section 2.6 and the equality (11) is obtained as follows. We have

$$\begin{aligned}
\|x_t - x_t(\vartheta)\|_{\nu_\varepsilon}^2 &= \int_0^{\nu_\varepsilon} [x_t - x_t(\vartheta)]^2 dt \\
&= \int_0^{\nu_\varepsilon} t^2 [S(0, x_0) - S(\vartheta, 0, x_0)]^2 dt (1 + o(1)).
\end{aligned}$$

Hence

$$\begin{aligned}
\bar{\vartheta}_{\nu_\varepsilon} &= \arg \inf_{\vartheta \in \Theta} \|x_t - x_t(\vartheta)\|_{\nu_\varepsilon}^2 \\
&= \arg \inf_{\vartheta \in \Theta} \frac{\nu_\varepsilon^3}{3} [S(0, x_0) - S(\vartheta, 0, x_0)]^2 (1 + o(1)) \longrightarrow \bar{\vartheta},
\end{aligned}$$

which yields (11).

Introduce the condition

$$\inf_{\bar{\vartheta}, \hat{\vartheta}} \sup_{0 \leq t \leq T} \left| \int_0^t \frac{\dot{S}(\bar{\vartheta}, s, x_s) [S(s, x_s) - S(\hat{\vartheta}, s, x_s)]}{\sigma(s, x_s)^2} ds \right| > 0.$$

It is easy to see that if this condition is fulfilled then  $\Delta_\varepsilon \rightarrow \infty$  and the test is consistent. Note that if this condition is not fulfilled then for all  $t \in [0, T]$  we have

$$\int_0^t \frac{\dot{S}(\bar{\vartheta}, s, x_s) \left[ S(s, x_s) - S(\hat{\vartheta}, s, x_s) \right]}{\sigma(s, x_s)^2} ds = 0$$

and this equality implies

$$\dot{S}(\bar{\vartheta}, t, x_t) \left[ S(t, x_t) - S(\hat{\vartheta}, t, x_t) \right] = 0, \quad 0 \leq t \leq T. \quad (12)$$

If  $\left| \dot{S}(\vartheta, t, x) \right| > 0$  for all  $\vartheta \in \Theta$  and almost all  $t \in [0, T]$  and almost all  $x \in K$  for any bounded region  $K \subset \mathcal{R}$ , then the proposed test is consistent against any fixed alternative.

An example of alternative invisible by this test can be constructed as follows. Suppose that the function  $S(\vartheta, t, x)$  does not depend on  $\vartheta$  for the values  $t \in [0, T/2]$  and the trend coefficient  $S(t, x_t)$  under alternative coincides with the function  $S(\vartheta^*, t, x_t)$  for  $t \in [T/2, T]$ . Then we have (12) in the situation, where the trend coefficients of diffusion process on the interval  $[0, T/2]$  can be different under alternative. Of course as we know that the trend coefficient under hypothesis does not depend on  $\vartheta$  on the interval  $[0, T/2]$ , then for this interval we can modify the test statistic.

**Example.** Suppose that the observed diffusion process under hypothesis has the stochastic differential

$$dX_t = \vartheta X_t dt + \varepsilon dW_t, \quad X_0 = x_0 > 0, \quad 0 \leq t \leq T,$$

where  $\vartheta \in \Theta$  and  $0 \notin \Theta$ . Then we have

$$I(\vartheta) = \frac{x_0^2(e^{2\vartheta T} - 1)}{2\vartheta}, \quad \hat{\vartheta}_\varepsilon = \frac{\int_0^T X_t dX_t}{\int_0^T X_t^2 dt}$$

and the statistic

$$\hat{U}_\varepsilon(t) = \frac{1}{\varepsilon x_0} \sqrt{\frac{2\hat{\vartheta}_\varepsilon T}{e^{2\hat{\vartheta}_\varepsilon T} - 1}} \int_0^t X_s \left[ dX_s - \hat{\vartheta}_\varepsilon X_s dt \right].$$

Here we have no problem of the definition of stochastic integral and this will always be the case for the models in which the trend coefficient depends linearly on the unknown parameter.

The test  $\hat{\psi}_\varepsilon = \mathbb{I}_{\{\Delta_\varepsilon > c_\alpha\}}$  with

$$\Delta_\varepsilon = \int_0^T \frac{\hat{U}_\varepsilon(t) X_t^2}{I(\hat{\vartheta}_\varepsilon) \sigma^2} dt \implies \int_0^1 B(\tau)^2 d\tau$$

is ADF.

## 2.2 Ergodic Diffusion Processes

Suppose that the observed diffusion process  $X^T = (X_t, 0 \leq t \leq T)$  satisfies the stochastic differential

$$dX_t = S(X_t) dt + \sigma(X_t) dW_t, \quad X_0, \quad 0 \leq t \leq T, \quad (13)$$

where the function  $\sigma(x)$  is known. The trend coefficient  $S(\cdot)$  is an unknown function and we have to test the following composite hypothesis:

$\mathcal{H}_0$  : The process  $X^T$  is the solution of equation

$$dX_t = S(\vartheta, X_t) dt + \sigma(X_t) dW_t, \quad X_0, \quad 0 \leq t \leq T, \quad \vartheta \in \Theta, \quad (14)$$

where  $S(\vartheta, x)$  is a known smooth function depending on unknown parameter  $\vartheta \in \Theta = (a, b)$ .

Introduce the regularity conditions.

$\mathcal{ES}$ . The function  $S(\vartheta, x)$  is locally bounded, the function  $\sigma(\cdot)^2 > 0$  is continuous and for some  $C > 0$  the condition

$$x S(\vartheta, x) + \sigma(x)^2 \leq C(1 + x^2)$$

holds.

By this condition the stochastic differential equation has a unique weak solution (see, e.g., [3]).

Let us denote by  $\mathcal{P}$  the class of locally bounded functions with polynomial majorants ( $p > 0$ )

$$\mathcal{P} = \{h(\cdot) : |h(y)| \leq C(1 + |y|^p)\}.$$

The next condition is

$\mathcal{A}_0$ . The functions  $S(\cdot), \sigma(\cdot)^{\pm 1} \in \mathcal{P}$  and

$$\overline{\lim}_{|y| \rightarrow \infty} \sup_{\vartheta \in \Theta} \operatorname{sgn}(y) \frac{S(\vartheta, y)}{\sigma(y)^2} < 0.$$

Note that if  $S(\vartheta, x)$  and  $\sigma(x)$  satisfy  $\mathcal{A}_0$ , then we have

$$V(\vartheta, x) = \int_0^x \exp \left\{ -2 \int_0^y \frac{S(\vartheta, z)}{\sigma(z)^2} dz \right\} dy \longrightarrow \pm \infty$$

as  $x \rightarrow \pm \infty$  and  $\sup_{\vartheta \in \Theta} G(\vartheta) < \infty$ , where

$$G(\vartheta) = \int_{-\infty}^{\infty} \sigma(y)^{-2} \exp \left\{ 2 \int_0^x \frac{S(\vartheta, y)}{\sigma(y)^2} dy \right\} dx$$

is normalizing constant.

By these conditions the stochastic process  $X^T$  is positive-recurrent (ergodic) with the density of the invariant law

$$f(\vartheta, x) = \frac{1}{G(\vartheta) \sigma(x)^2} \exp \left\{ 2 \int_0^x \frac{S(\vartheta, y)}{\sigma(y)^2} dy \right\}.$$

Let us introduce further regularity conditions.

$\mathcal{R}_e$ . The function  $S(\vartheta, x)$  has two continuous derivatives

$$\dot{S}(\vartheta, x), \ddot{S}(\vartheta, x) \in \mathcal{P}.$$

and

$\mathcal{I}_e$ . For any  $\nu > 0$

$$\inf_{\vartheta_0 \in \Theta} \inf_{|\vartheta - \vartheta_0| > \nu} \int_{-\infty}^{\infty} \left[ \frac{S(\vartheta, x) - S(\vartheta_0, x)}{\sigma(x)} \right]^2 f(\vartheta_0, x) dx > 0.$$

The likelihood ratio function is

$$L(\vartheta, X^T) = \exp \left\{ \int_0^T \frac{S(\vartheta, X_t)}{\sigma(X_t)^2} dX_t - \int_0^T \frac{S(\vartheta, X_t)^2}{2\sigma(X_t)^2} dt \right\}.$$

Under the regularity conditions assumed above, the MLE  $\hat{\vartheta}_T$  admits the representation

$$\sqrt{T}(\hat{\vartheta}_T - \vartheta) = \frac{1}{I(\vartheta)\sqrt{T}} \int_0^T \frac{\dot{S}(\vartheta, X_t)}{\sigma(X_t)} dW_t + o(1).$$

Here  $I(\vartheta)$  is the Fisher information

$$I(\vartheta) = \int_{-\infty}^{\infty} \left( \frac{\dot{S}(\vartheta, x)}{\sigma(x)} \right)^2 f(\vartheta, x) dx > 0.$$

The proof can be found in [12].

The score-function is

$$\frac{\partial \ln L(\vartheta, X^T)}{\partial \vartheta} = \int_0^T \frac{\dot{S}(\vartheta, X_t)}{\sigma(X_t)^2} [dX_t - S(\vartheta, X_t) dt]$$

and we define the normalized score-function:

$$U_T(\vartheta, X^T) = \varphi_T(\vartheta) \int_0^T \frac{\dot{S}(\vartheta, X_t)}{\sigma(X_t)^2} [dX_t - S(\vartheta, X_t) dt] \implies \xi,$$



where  $\varphi_T(\vartheta) = [T\mathcal{I}(\vartheta)]^{-1/2}$ . The limit random variable  $\xi$  can be written as the following integral

$$\xi = \int_{-\infty}^{\infty} \frac{\dot{S}(\vartheta, y) \sqrt{f(\vartheta, y)}}{\sqrt{\mathcal{I}(\vartheta)} \sigma(y)} dw(y) \sim \mathcal{N}(0, 1),$$

where  $w(\cdot)$  is two-sided Wiener process.

Let us introduce the slightly modified score-function process

$$U_T(x, \vartheta, X^T) = \varphi_T(\vartheta) \int_0^T \frac{\dot{S}(\vartheta, X_t)}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} [dX_t - S(\vartheta, X_t) dt], \quad x \in \mathcal{R},$$

and (formally) the statistic

$$\hat{U}_T(x) = U_T(x, \hat{\vartheta}_T, X^T).$$

Note that with  $\mathbf{P}_{\vartheta_0}$  probability 1 we have the equality  $U_T(\infty, \vartheta_0, X^T) = U_T(\vartheta_0, X^T)$ . The asymptotic behaviour of this statistic can be explained as follows (again, formally).

$$\begin{aligned} \hat{U}_T(x) &= \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t)}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} [dX_t - S(\hat{\vartheta}_T, X_t) dt] \\ &= \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t)}{\sigma(X_t)} \mathbb{I}_{\{X_t < x\}} dW_t \\ &\quad + \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t) [S(\vartheta_0, X_t) - S(\hat{\vartheta}_T, X_t)]}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dt \\ &= \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t)}{\sigma(X_t)} \mathbb{I}_{\{X_t < x\}} dW_t \\ &\quad - \frac{\hat{\vartheta}_T - \vartheta_0}{\varphi_T(\hat{\vartheta}_T)} \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t) \dot{S}(\vartheta_0, X_t)}{T\mathcal{I}(\hat{\vartheta}_T) \sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dt \\ &= \varphi_T(\vartheta_0) \int_0^T \frac{\dot{S}(\vartheta_0, X_t)}{\sigma(X_t)} \mathbb{I}_{\{X_t < x\}} dW_t \\ &\quad - \int_0^T \frac{\dot{S}(\vartheta_0, X_t)}{\sqrt{T\mathcal{I}(\vartheta_0)} \sigma(X_t)} dW_t \int_0^T \frac{\dot{S}(\vartheta_0, X_t)^2}{T\mathcal{I}(\vartheta_0) \sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dt + o(1). \end{aligned} \tag{15}$$

Here  $\vartheta_0$  is the true value of the parameter. These integrals have the following

limits

$$\begin{aligned} \frac{1}{\sqrt{T\mathbb{I}(\vartheta_0)}} \int_0^T \frac{\dot{S}(\vartheta_0, X_t)}{\sigma(X_t)} \mathbb{I}_{\{X_t < x\}} dW_t &\Longrightarrow \int_{-\infty}^x \frac{\dot{S}(\vartheta_0, y) \sqrt{f(\vartheta_0, y)}}{\sqrt{\mathbb{I}(\vartheta_0)} \sigma(y)} dw(y), \\ \frac{1}{T\mathbb{I}(\vartheta_0)} \int_0^T \frac{\dot{S}(\vartheta_0, X_t)^2}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dt &\longrightarrow \int_{-\infty}^x \frac{\dot{S}(\vartheta_0, y)^2 f(\vartheta_0, y)}{\mathbb{I}(\vartheta_0) \sigma(y)^2} dy. \end{aligned}$$

Let us denote

$$\tau = \int_{-\infty}^x \frac{\dot{S}(\vartheta_0, y)^2 f(\vartheta_0, y)}{\mathbb{I}(\vartheta_0) \sigma(y)^2} dy, \quad 0 \leq \tau \leq 1.$$

Then we have the convergence

$$\hat{U}_T(x) \Longrightarrow W(\tau) - W(1) \tau = B(\tau), \quad 0 \leq \tau \leq 1.$$

This limit suggests the construction of the statistic

$$\Delta_T = \int_{-\infty}^{\infty} \frac{\hat{U}_T(x)^2 \dot{S}(\hat{\vartheta}_T, x)^2}{\mathbb{I}(\hat{\vartheta}_T) \sigma(x)^2} dF(\hat{\vartheta}_T, x)$$

and the test

$$\hat{\psi}_T = \mathbb{I}_{\{\Delta_T > c_\alpha\}}, \quad \mathbf{P}(\Delta > c_\alpha) = \alpha.$$

Note that

$$\tau_T = \int_{-\infty}^x \frac{\dot{S}(\hat{\vartheta}_T, x)^2}{\mathbb{I}(\hat{\vartheta}_T) \sigma(x)^2} dF(\hat{\vartheta}_T, x) \longrightarrow \tau.$$

Hence if we verify that  $\Delta_T \Rightarrow \Delta$ , then the test  $\hat{\psi}_T \in \mathcal{K}_\alpha$  and is ADF.

We have the same problem with the definition of the stochastic integral

$$\int_0^T \frac{\dot{S}(\hat{\vartheta}_t, X_t)}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dX_t$$

as in (6) and we propose two approaches. In the first one we replace it by the ordinary integral using the Itô formula as it was done above and in the second approach we propose using a preliminary consistent estimator of the parameter  $\vartheta$ .

*First approach.* Introduce the function,

$$H_T(\vartheta, x, z) = \int_{X_0}^z \frac{\dot{S}(\vartheta, y)}{\sigma(y)^2} \phi_T(x - y) dy, \quad H(\vartheta, x, z) = \int_{X_0}^z \frac{\dot{S}(\vartheta, y)}{\sigma(y)^2} \mathbb{I}_{\{y < x\}} dy$$

where  $\phi_T(x - y)$  is a “smooth approximation” of the indicator function  $\mathbb{I}_{\{y < x\}}$ . For example,  $\phi_T(x - y) = \phi\left(\frac{x-y}{d_T}\right)$ , where

$$\phi(z) = a^{-1} \int_{-\infty}^z e^{\frac{v^2}{v^2-1}} \mathbb{I}_{\{|v|<1\}} dv, \quad a = \int_{-1}^1 e^{\frac{v^2}{v^2-1}} \mathbb{I}_{\{|v|<1\}} dv$$

and  $d_T \rightarrow 0$ .

We write

$$\begin{aligned} \int_0^T \frac{\dot{S}(\vartheta, X_t)}{\sigma(X_t)^2} \phi_T(x - X_t) dX_t &= H_T(\vartheta, x, X_T) \\ &\quad - \frac{1}{2} \int_0^T \sigma(X_s)^2 (H_T)''_{z,z}(\vartheta, x, X_s) ds. \end{aligned}$$

Then we use the representation of the modified score-function process  $\tilde{U}_T(x)$  (we replaced the indicator function by its smooth approximation)

$$\begin{aligned} \tilde{U}_T(x) &= \varphi_T(\hat{\vartheta}_T) H_T(\hat{\vartheta}_T, x, X_T) - \frac{\varphi_T(\hat{\vartheta}_T)}{2} \int_0^T \sigma(X_s)^2 (H_T)''_{z,z}(\hat{\vartheta}_T, x, X_s) ds \\ &\quad - \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t) S(\vartheta_0, X_t)}{\sigma(X_t)^2} \phi_T(x - X_s) dt \\ &\quad + \varphi_T(\hat{\vartheta}_T) \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t) [S(\vartheta_0, X_t) - S(\hat{\vartheta}_T, X_t)]}{\sigma(X_t)^2} \phi_T(x - X_s) dt \\ &= J_T(\hat{\vartheta}_T, x) - K_T(\hat{\vartheta}_T, x). \end{aligned}$$

Direct but cumbersome calculations give the limits

$$\begin{aligned} J_T(\hat{\vartheta}_T, x) &\Rightarrow \int_{-\infty}^x \frac{\dot{S}(\vartheta_0, y) \sqrt{f(\vartheta_0, y)}}{\sqrt{I(\vartheta_0)} \sigma(y)} dw(y), \\ K_T(\hat{\vartheta}_T, x) &= \frac{(\hat{\vartheta}_T - \vartheta)}{\varphi_T(\hat{\vartheta}_T)} \int_0^T \frac{\dot{S}(\hat{\vartheta}_T, X_t) \dot{S}(\tilde{\vartheta}_T, X_t)}{TI(\hat{\vartheta}_T) \sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} dt (1 + o(1)) \\ &\Rightarrow \int_{-\infty}^{\infty} \frac{\dot{S}(\vartheta_0, y) \sqrt{f(\vartheta_0, y)}}{\sqrt{I(\vartheta_0)} \sigma(y)} dw(y) \int_{-\infty}^x \frac{\dot{S}(\vartheta_0, y) \sqrt{f(\vartheta_0, y)}}{\sqrt{I(\vartheta_0)} \sigma(y)} dy. \end{aligned}$$

Thus we have the following result.

**Proposition 3** Suppose that the conditions of regularity are fulfilled, then the test  $\psi_T = \mathbb{I}_{\{\tilde{\Delta}_T > c_\alpha\}}$  with

$$\tilde{\Delta}_T = \int_{-\infty}^{\infty} \frac{\tilde{U}_T(x)^2 \dot{S}(\hat{\vartheta}_T, x)^2}{\mathbb{I}(\hat{\vartheta}_T) \sigma(x)^2} dF(\hat{\vartheta}_T, x)$$

is ADF and belongs to  $\mathcal{K}_\alpha$ .

*The second approach.* Let us introduce a consistent preliminary estimator  $\bar{\vartheta}_{\sqrt{T}}$  constructed using the first  $X_t, 0 \leq t \leq \sqrt{T}$  observations. For example, the method of moments estimator can be used (see conditions of consistency in [12], Section 2.4). The corresponding statistic is

$$V_T(x) = \varphi_T(\bar{\vartheta}_{\sqrt{T}}) \int_{\sqrt{T}}^T \frac{\dot{S}(\bar{\vartheta}_{\sqrt{T}}, X_t)}{\sigma(X_t)^2} \mathbb{I}_{\{X_t < x\}} [dX_t - S(\hat{\vartheta}_T, X_t) dt].$$

The stochastic integral is well defined and its limit can be obtained calculations, similar to (15).

**Proposition 4** Suppose that the conditions of regularity are fulfilled and the preliminary estimator  $\bar{\vartheta}_{\sqrt{T}}$  is consistent, then the test  $\psi_T = \mathbb{I}_{\{\tilde{\Delta}_T > c_\alpha\}}$  with

$$\tilde{\Delta}_T = \int_{-\infty}^{\infty} \frac{V_T(x)^2 \dot{S}(\bar{\vartheta}_T, x)^2}{\mathbb{I}(\bar{\vartheta}_T) \sigma(x)^2} dF(\bar{\vartheta}_T, x)$$

is ADF and belongs to  $\mathcal{K}_\alpha$ .

The condition of the consistency is

$$\inf_{\bar{\vartheta}, \hat{\vartheta}} \int_{-\infty}^{\infty} \frac{M(\bar{\vartheta}, \hat{\vartheta}, x)^2 \dot{S}(\bar{\vartheta}, x)^2}{\sigma(x)^2} f(x) dx > 0,$$

where

$$M(\bar{\vartheta}, \hat{\vartheta}, x) = \int_{-\infty}^x \frac{\dot{S}(\bar{\vartheta}, y) [S(y) - S(\hat{\vartheta}, y)]}{\sigma(y)^2} f(y) dy.$$

## 2.3 Periodic Poisson Processes.

The last observations model is a periodic Poisson process

$$X^n = (X_t, 0 \leq t \leq T = n\tau_*)$$

of known period  $\tau_* > 0$ . For  $0 \leq s < t$  and  $k = 0, 1, 2, \dots$

$$\mathbf{P}(X_t - X_s = k) = \frac{[\Lambda(t) - \Lambda(s)]^k}{k!} \exp\{-\Lambda(t) + \Lambda(s)\}.$$

The mean  $\Lambda(t)$  and intensity function  $\lambda(t)$  satisfy the relations

$$\Lambda(t) = \mathbf{E}X_t, \quad \Lambda(t) = \int_0^t \lambda(s) \, ds$$

and  $\lambda(t + k\tau_*) = \lambda(t)$ .

We observe a trajectory  $X^n$  of the Poisson process of intensity function  $\lambda(\cdot)$  and we have to test the hypothesis

$\mathcal{H}_0$  : The intensity function  $\lambda(t) = \lambda(\vartheta, t)$ ,  $\vartheta \in \Theta = (a, b)$ .

Here  $\lambda(\vartheta, \cdot)$  is some known function satisfying the following conditions of regularity.

The intensity function  $\lambda(\vartheta, \cdot)$  is twice continuously differentiable w.r.t.  $\vartheta$ , strictly positive and the identifiability condition holds: for any  $\nu > 0$

$$\inf_{\vartheta_0 \in \Theta} \inf_{|\vartheta - \vartheta_0| > \nu} \int_0^{\tau_*} \left[ \sqrt{\lambda(\vartheta, s)} - \sqrt{\lambda(\vartheta_0, s)} \right]^2 \, ds > 0.$$

The likelihood ratio function is

$$L(\vartheta, X^n) = \exp \left\{ \sum_{j=1}^n \int_0^{\tau_*} \ln \lambda(\vartheta, t) \, dX_j(t) - n \int_0^{\tau_*} [\lambda(\vartheta, t) - 1] \, dt \right\}$$

and the MLE  $\hat{\vartheta}_n$  is defined by the equation like (4). Then the MLE admits the representation

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) = \frac{1}{I(\vartheta) \sqrt{n}} \sum_{j=1}^n \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} \, d[X_j(s) - \lambda(\vartheta, s) \, ds] + o(1).$$

Here  $X_j(s) = X_{(j-1)\tau_*+s} - X_{(j-1)\tau_*}$ ,  $0 \leq s \leq \tau_*$ ,  $j = 1, 2, \dots, n$  and  $I(\vartheta)$  is the Fisher information

$$I(\vartheta) = \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, s)^2}{\lambda(\vartheta, s)} \, ds > 0.$$

The proof can be found in [11].

The score-function for this process is

$$\frac{\partial \ln L(\vartheta, X^n)}{\partial \vartheta} = \sum_{j=1}^n \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, t)}{\lambda(\vartheta, t)} [dX_j(t) - \lambda(\vartheta, t) \, dt]$$

and we define the normalized score-function process

$$U_n(t, \vartheta, X^n) = \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} [dX_j(s) - \lambda(\vartheta, s) ds].$$

We construct the GoF test with the help of the statistic

$$\hat{U}_n(t) = U_n(t, \hat{\vartheta}_n, X^n)$$

Its formal expansion provides us with the following expressions (we put below  $\pi_j(s) = X_j(s) - \Lambda(\vartheta, s)$ )

$$\begin{aligned} \hat{U}_n(t) &= \frac{1}{\sqrt{I(\hat{\vartheta}_n)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\hat{\vartheta}_n, s)}{\lambda(\hat{\vartheta}_n, s)} [dX_j(s) - \lambda(\hat{\vartheta}_n, s) ds] \\ &= \frac{1}{\sqrt{I(\hat{\vartheta}_n)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\hat{\vartheta}_n, s)}{\lambda(\hat{\vartheta}_n, s)} [dX_j(s) - \lambda(\vartheta, s) ds] \\ &\quad + \frac{1}{\sqrt{I(\hat{\vartheta}_n)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\hat{\vartheta}_n, s)}{\lambda(\hat{\vartheta}_n, s)} [\lambda(\vartheta, s) - \lambda(\hat{\vartheta}_n, s)] ds \\ &= \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} d\pi_j(s) - \frac{\sqrt{n}(\hat{\vartheta}_n - \vartheta)}{\sqrt{I(\vartheta)}} \int_0^t \frac{\dot{\lambda}(\vartheta, s)^2}{\lambda(\vartheta, s)} ds + o(1) \\ &= \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} d\pi_j(s) \\ &\quad - \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} d\pi_j(s) \int_0^t \frac{\dot{\lambda}(\vartheta, s)^2}{I(\vartheta) \lambda(\vartheta, s)} ds + o(1). \end{aligned}$$

By the central limit theorem we have the convergence in distribution

$$\begin{aligned} \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^t \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} d\pi_j(s) &\Rightarrow \frac{1}{\sqrt{I(\vartheta)}} \int_0^t \frac{\dot{\lambda}(\vartheta, s)}{\sqrt{\lambda(\vartheta, s)}} dW_s, \\ \frac{1}{\sqrt{I(\vartheta)}n} \sum_{j=1}^n \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, s)}{\lambda(\vartheta, s)} d\pi_j(s) &\Rightarrow \frac{1}{\sqrt{I(\vartheta)}} \int_0^{\tau_*} \frac{\dot{\lambda}(\vartheta, s)}{\sqrt{\lambda(\vartheta, s)}} dW_s, \end{aligned}$$

where  $W_t, 0 \leq t \leq \tau_*$  is some Wiener process. Therefore, if we put

$$\tau = \int_0^t \frac{\dot{\lambda}(\vartheta, s)^2}{I(\vartheta) \lambda(\vartheta, s)} ds, \quad 0 \leq \tau \leq 1,$$

then once again we obtain the convergence

$$\hat{U}_n(t) \implies W(\tau) - W(1)\tau = B(\tau), \quad 0 \leq \tau \leq 1.$$

We can consider two approaches as before, but we present here the second (more simple) construction of the test. Let us take any consistent estimator  $\bar{\vartheta}_N$  of the parameter  $\vartheta$  constructed by the first  $N = \lfloor \sqrt{n} \rfloor$  observations  $X^N = (X_1, \dots, X_N)$ . Then we set

$$V_n(t) = \frac{1}{\sqrt{I(\bar{\vartheta}_N)n}} \sum_{j=N+1}^n \int_0^t \frac{\dot{\lambda}(\bar{\vartheta}_N, s)}{\lambda(\bar{\vartheta}_N, s)} \left[ dX_j(s) - \lambda(\bar{\vartheta}_N, s) ds \right].$$

The estimator  $\bar{\vartheta}_N$  and the observations  $X_{N+1}^n = (X_{N+1}, \dots, X_n)$  are independent and the stochastic integral with respect to the Poisson process is well defined (see Liptser, Shirayev [18], Section 18.4).

**Proposition 5** *Let the conditions of regularity be fulfilled, then the test  $\tilde{\psi}_n = \mathbb{I}_{\{\Delta_n > c_\alpha\}}$  with*

$$\Delta_n = \int_0^{\tau_*} \frac{V_n(t)^2 \dot{\lambda}(\bar{\vartheta}_N, s)^2}{I(\bar{\vartheta}_N) \lambda(\bar{\vartheta}_N, s)} ds.$$

*is ADF and belongs to  $\mathcal{K}_\alpha$ .*

To prove this proposition we have to verify the convergence

$$\Delta_n \implies \Delta = \int_0^1 B(\tau)^2 d\tau$$

under hypothesis  $\mathcal{H}_0$

**Example.** Suppose that the intensity function under hypothesis  $\mathcal{H}_0$  is

$$\lambda(\vartheta, t) = \vartheta h(t) + \lambda_0, \quad 0 \leq t \leq \tau_*,$$

where  $\vartheta \in \Theta = (a, b)$ ,  $a > 0$  and the function  $h(t) > 0$ .

Then we can take as preliminary estimator the minimum distance estimator

$$\begin{aligned} \bar{\vartheta}_N &= \arg \inf_{\vartheta \in \Theta} \int_0^{\tau_*} \left[ \hat{\Lambda}_N(t) - \vartheta H(t) - \lambda_0 t \right]^2 dt \\ &= \frac{\int_0^{\tau_*} \left[ \hat{\Lambda}_N(t) - \lambda_0 t \right] H(t) dt}{\int_0^{\tau_*} H(t)^2 dt}. \end{aligned}$$

Here

$$\hat{\Lambda}_N(t) = \frac{1}{N} \sum_{j=1}^N X_j(t), \quad H(t) = \int_0^t h(s) \, ds.$$

This is an unbiased, consistent and asymptotically normal estimator of the parameter  $\vartheta$ .

The score-function process  $V_n(\cdot)$  and the test statistics  $\tilde{\Delta}_n$  are

$$V_n(t) = \frac{1}{\sqrt{I(\vartheta_N)n}} \sum_{j=N+1}^n \int_0^t \frac{h(s)}{\vartheta_N h(s) + \lambda_0} [dX_j(s) - [\bar{\vartheta}_N h(s) + \lambda_0] \, ds],$$

$$\tilde{\Delta}_n = \int_0^{\tau^*} \frac{V_n(t)^2 h(s)^2}{I(\vartheta_N) [\bar{\vartheta}_N h(s) + \lambda_0]} \, ds, \quad I(\vartheta) = \int_0^{\tau^*} \frac{h(t)^2}{\vartheta h(t) + \lambda_0} \, dt,$$

respectively.

### 3 Other tests and models

#### 3.1 Other tests

The statistics  $U_\varepsilon(\cdot)$ ,  $U_T(\cdot)$  and  $U_n(\cdot)$  can be used for construction of the ADF GoF tests of Kolmogorov-Smirnov type. For example, the following convergence

$$\Delta_\varepsilon^* = \sup_{\nu_\varepsilon \leq t \leq T} |V_\varepsilon(t)| \implies \sup_{0 \leq \tau \leq 1} |B(\tau)| = \Delta^*$$

can be easily proved. Hence the test

$$\psi_\varepsilon^* = \mathbb{1}_{\{\Delta_\varepsilon^* > d_\alpha\}}, \quad \mathbf{P}(\Delta^* > d_\alpha) = \alpha$$

belongs to  $\mathcal{K}_\alpha$  and is ADF. Of course similar tests can be constructed in the cases of observations of the ergodic diffusion and inhomogeneous Poisson processes as well.

#### 3.2 Nonlinear AR process

Suppose that the observations  $X^n = (X_0, X_1, \dots, X_n)$  satisfy the relation

$$X_j = S(X_{j-1}) + \varepsilon_j, \quad j = 1, \dots, n$$



and we have to test a parametric hypothesis

$$\mathcal{H}_0 \quad : \quad S(x) = S(\vartheta, x), \quad \vartheta \in \Theta = (a, b).$$

Here  $S(\vartheta, x)$  is some known function and  $\vartheta$  is the unknown parameter. The random variables  $\varepsilon_1, \dots, \varepsilon_j$  are i.i.d. with the known density function  $f(x)$ .

The functions  $S(\vartheta, x)$  and  $f(x) > 0$  are such that the time series  $(X_j)_{j \geq 1}$  has ergodic properties with the density of invariant law  $\varphi(\vartheta, x)$  for all  $\vartheta \in \Theta$ , i.e., for any function  $h(\cdot)$  such that  $\mathbf{E}_\vartheta |h(\xi)| < \infty$  (here  $\xi \sim \varphi(\vartheta, \cdot)$ ) we have the law of large numbers

$$\frac{1}{n} \sum_{j=1}^n h(X_j) \longrightarrow \mathbf{E}_\vartheta h(\xi).$$

Moreover we suppose that the tails of  $\varphi(\vartheta, x)$  decrease sufficiently fast

$$\varphi(\vartheta, x) \leq \frac{C}{|x|^{1+\gamma}} \quad (16)$$

with some positive constants  $\gamma$  and  $C$ , which do not depend on  $\vartheta$ . The log-density function  $\ell(x) = \ln f(x)$  has three continuous bounded derivatives  $\ell'(x)$ ,  $\ell''(x)$ ,  $\ell'''(x)$  and the function  $S(\vartheta, x)$  has two continuous bounded derivatives  $\dot{S}(\vartheta, x)$ ,  $\ddot{S}(\vartheta, x)$  w.r.t.  $\vartheta$ .

The log-likelihood function is

$$L(\vartheta, X^n) = \ln \varphi(\vartheta, X_0) + \sum_{j=1}^n \ln f(X_j - S(\vartheta, X_{j-1})), \quad \vartheta \in (a, b).$$

We suppose that the initial value  $X_0$  has invariant density function  $\varphi(\vartheta, x)$  and therefore the time series  $(X_j)_{j \geq 0}$  is stationary.

The Score-function is

$$U_n(\vartheta, X^n) = - \sum_{j=1}^n \ell'(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1}).$$

Also we assume that the regularity conditions are fulfilled so that the MLE  $\hat{\vartheta}_n$  is consistent and admits the representation

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) = \frac{-1}{\mathbf{I}(\vartheta) \sqrt{n}} \sum_{j=1}^n \ell'(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1}) + o(1), \quad (17)$$

where the Fisher information

$$\mathbf{I}(\vartheta) = \mathbb{E}_\vartheta \left[ \ell'(\varepsilon_1) \dot{S}(\vartheta, \xi) \right]^2 = \mathbf{I}_f \mathbf{I}_\vartheta.$$

Here we denoted  $\mathbb{E}_\vartheta$  the expectation related to the couple of independent random variables  $(\varepsilon, \xi)$ , i.e.,

$$I_f = \mathbf{E} \ell'(\varepsilon)^2 = \int_{-\infty}^{\infty} \frac{f'(x)^2}{f(x)} dx, \quad I_\vartheta = \mathbf{E}_\vartheta \dot{S}(\vartheta, \xi)^2 = \int_{-\infty}^{\infty} \dot{S}(\vartheta, x)^2 \varphi(\vartheta, x) dx.$$

Note that from this representation and the central limit theorem it follows that the MLE is asymptotically normal (see, e.g., [6])

$$\hat{u}_n = \sqrt{n} (\hat{\vartheta}_n - \vartheta) \implies \mathcal{N}(0, I(\vartheta)^{-1}).$$

Introduce the normalized score-function process

$$U_n(x, \vartheta, X^n) = \frac{-1}{\sqrt{I(\vartheta)n}} \sum_{j=1}^n \ell'(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1}) \mathbb{I}_{\{X_{j-1} < x\}}$$

and the corresponding statistics

$$\hat{U}_n(x) = \frac{-1}{\sqrt{I(\hat{\vartheta}_n)n}} \sum_{j=1}^n \ell'(X_j - S(\hat{\vartheta}_n, X_{j-1})) \dot{S}(\hat{\vartheta}_n, X_{j-1}) \mathbb{I}_{\{X_{j-1} < x\}}.$$

Using the expansion at the vicinity of the true value  $\vartheta$  we can write

$$\begin{aligned} \hat{U}_n(x) &= \frac{-1}{\sqrt{I(\vartheta)n}} \sum_{j=1}^n [\ell'(X_j - S(\vartheta, X_{j-1})) \\ &\quad - \frac{\hat{u}_n}{\sqrt{n}} \ell''(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1})] \dot{S}(\vartheta, X_{j-1}) \mathbb{I}_{\{X_{j-1} < x\}} + o(1) \\ &= \frac{-1}{\sqrt{I(\vartheta)n}} \sum_{j=1}^n \ell'(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1}) \mathbb{I}_{\{X_{j-1} < x\}} \\ &\quad + \frac{\hat{u}_n \sqrt{I(\vartheta)}}{I(\vartheta)n} \sum_{j=1}^n \ell''(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1})^2 \mathbb{I}_{\{X_{j-1} < x\}} + o(1). \end{aligned}$$

The standard arguments allow us to write

$$\begin{aligned} &\frac{-1}{I(\vartheta)n} \sum_{j=1}^n \ell''(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1})^2 \mathbb{I}_{\{X_{j-1} < x\}} \\ &= \frac{-1}{I(\vartheta)n} \sum_{j=1}^n \ell''(\varepsilon_j) \dot{S}(\vartheta, X_{j-1})^2 \mathbb{I}_{\{X_{j-1} < x\}} \longrightarrow \frac{1}{I_\vartheta} \int_{-\infty}^x \dot{S}(\vartheta, y)^2 \varphi(\vartheta, y) dy. \end{aligned}$$

Recall, that

$$\mathbf{E} \ell''(\varepsilon) = \mathbf{E} \left( \frac{f''(\varepsilon) f(\varepsilon) - f'(\varepsilon)^2}{f(\varepsilon)^2} \right) = -\mathbf{E} \left( \frac{f'(\varepsilon)}{f(\varepsilon)} \right)^2 = -I_f$$

because

$$\mathbf{E} \left( \frac{f''(\varepsilon)}{f(\varepsilon)} \right) = \int_{-\infty}^{\infty} f''(y) dy = 0.$$

Let us denote

$$W_n(x) = \frac{-1}{\sqrt{I(\vartheta)n}} \sum_{j=1}^n \ell'(X_j - S(\vartheta, X_{j-1})) \dot{S}(\vartheta, X_{j-1}) \mathbb{1}_{\{X_{j-1} < x\}}.$$

We have

$$\begin{aligned} & \mathbf{E}_{\vartheta} W_n(x) W_n(y) \\ &= \frac{1}{I(\vartheta)n} \sum_{j=1}^n \sum_{i=1}^n \mathbf{E}_{\vartheta} \ell'(\varepsilon_j) \ell'(\varepsilon_i) \dot{S}(\vartheta, X_{j-1}) \dot{S}(\vartheta, X_{i-1}) \mathbb{1}_{\{X_{j-1} < x\}} \mathbb{1}_{\{X_{i-1} < y\}} \\ &= \frac{1}{I(\vartheta)} \mathbb{E}_{\vartheta} \ell'(\varepsilon_1)^2 \dot{S}(\vartheta, \xi)^2 \mathbb{1}_{\{\xi < x \wedge y\}} \\ &= \min \left( I_{\vartheta}^{-1} \int_{-\infty}^x \dot{S}(\vartheta, z)^2 \varphi(\vartheta, z) dz, I_{\vartheta}^{-1} \int_{-\infty}^y \dot{S}(\vartheta, z)^2 \varphi(\vartheta, z) dz \right) \\ &= \min(\tau_x, \tau_y), \quad 0 \leq \tau_x = I_{\vartheta}^{-1} \int_{-\infty}^x \dot{S}(\vartheta, z)^2 \varphi(\vartheta, z) dz \leq 1. \end{aligned}$$

It can be shown that by the central limit theorem the finite-dimensional distributions of the random function  $W_n(x), x \in \mathbb{R}$  converge to the finite-dimensional distributions of the Wiener process  $W(\tau_x), x \in \mathbb{R}$ . Moreover the following estimate holds

$$\mathbf{E}_{\vartheta} |W_n(x) - W_n(y)|^2 \leq C |x - y|. \quad (18)$$

We have similar convergence for the MLE due to the representation (17)

$$\sqrt{nI(\vartheta)} (\hat{\vartheta}_n - \vartheta) = W_n(\infty) + o(1) \implies W(1)$$

with the same Wiener process, i.e., we have the joint asymptotic normality of  $W_n(\cdot)$  and  $\hat{u}_n$ . Therefore the random functions  $\hat{U}_n(x)$  have the corresponding limit

$$\hat{U}_n(x) \implies W(\tau_x) - W(1) \tau_x = B(\tau_x)$$

and again, we obtain the Brownian bridge  $B(\tau), 0 \leq \tau \leq 1$ .

Let us introduce the statistics

$$\Delta_n = \int_{-\infty}^{\infty} \frac{\hat{U}_n(x)^2 \dot{S}(\hat{\vartheta}_n, x)^2}{I_{\hat{\vartheta}_n}} \varphi(\hat{\vartheta}_n, x) dx.$$

The convergence of finite-dimensional distributions, the estimate (18) and the condition (16) allow us to verify the convergence

$$\Delta_n \Rightarrow \int_{-\infty}^{\infty} \frac{B(\tau_x)^2 \dot{S}(\vartheta, x)^2}{I_{\vartheta}} \varphi(\vartheta, x) dx = \int_0^1 B(\tau)^2 d\tau.$$

Therefore we have the following result.

**Proposition 6** *The test  $\hat{\psi}_n = \mathbb{I}_{\{\Delta_n > c_\alpha\}}$  is ADF and belongs to the class  $\mathcal{K}_\alpha$ .*

**Example.** Suppose that the observed time series  $(X_j)_{j \geq 1}$  under the hypothesis  $\mathcal{H}_0$  is linear AR

$$X_j = \vartheta X_{j-1} + \varepsilon_j, \quad j = 1, \dots, n,$$

where  $\vartheta \in \Theta = (-1, 1)$  and  $(\varepsilon_j)_{j \geq 1}$  are i.i.d.  $\mathcal{N}(0, \sigma^2)$  r.v.'s. Then the aforementioned conditions are satisfied with the density of invariant law

$$\varphi(\vartheta, x) \sim \mathcal{N}\left(0, \frac{\sigma^2}{1 - \vartheta^2}\right)$$

and we assume that  $X_0 \sim \varphi(\vartheta, x)$ .

The derivative  $\dot{S}(\vartheta, x) = x$  is not bounded, but the tails of  $\varphi(\vartheta, x)$  are exponentially decreasing and the proof of the convergence given above remains valid.

The score-function process is

$$U_n(x, \vartheta, X^n) = \frac{1}{\sqrt{n(1 - \vartheta^2)}} \sum_{j=1}^n (X_j - \vartheta X_{j-1}) X_{j-1} \mathbb{I}_{\{X_{j-1} < x\}},$$

because

$$I(\vartheta) = I_f I_{\vartheta} = \frac{1}{\sigma^2} \frac{\sigma^2}{1 - \vartheta^2} = \frac{1}{1 - \vartheta^2}$$

and we put

$$\hat{U}_n(x) = U_n(x, \hat{\vartheta}_n, X^n), \quad \hat{\vartheta}_n = \frac{\sum_{j=1}^n X_j X_{j-1}}{\sum_{j=1}^n X_{j-1}^2}.$$

Introduce the statistic

$$\Delta_n = \frac{1 - \hat{\vartheta}_n^2}{\sigma^2} \int_{-\infty}^{\infty} \hat{U}_n(x)^2 x^2 \varphi(\hat{\vartheta}_n, x) dx.$$

As it follows from the Proposition 6

$$\Delta_n \Rightarrow \int_0^1 B(\tau)^2 d\tau$$

and the test  $\hat{\psi}_n = \mathbb{I}_{\{\Delta_n > c_\alpha\}}$  is ADF and belongs to  $\mathcal{K}_\alpha$ .

### 3.3 The case of i.i.d. observations

Let us see what happens if we apply the same approach in the case of i.i.d. observations  $X^n = (X_1, \dots, X_n)$ , where  $X_j$  has the density function  $f(x)$ . Suppose that we have a parametric hypothesis

$$\mathcal{H}_0, \quad : \quad f(x) = f(\vartheta, x), \quad \vartheta \in \Theta = (a, b).$$

Here  $f(\vartheta, x)$  is some known density function satisfying the regularity conditions, which validate the calculations below.

The normalized score-function statistic is

$$\begin{aligned} U_n(\vartheta, X^n) &= \frac{1}{\sqrt{I(\vartheta)} n} \sum_{j=1}^n \dot{\ell}(\vartheta, X_j) = \frac{\sqrt{n}}{\sqrt{I(\vartheta)}} \int_{-\infty}^{\infty} \dot{\ell}(\vartheta, y) d\hat{F}_n(y) \\ &= \frac{\sqrt{n}}{\sqrt{I(\vartheta)}} \int_{-\infty}^{\infty} \dot{\ell}(\vartheta, y) \left[ d\hat{F}_n(y) - f(\vartheta, y) dy \right], \end{aligned}$$

where  $\ell(\vartheta, y) = \ln f(\vartheta, y)$ ,  $I(\vartheta)$  is the Fisher information and we used the equality

$$\int_{-\infty}^{\infty} \dot{\ell}(\vartheta, y) f(\vartheta, y) dy = 0.$$

Introduce the score-function process

$$U_n(\vartheta, x, X^n) = \frac{\sqrt{n}}{\sqrt{I(\vartheta)}} \int_{-\infty}^x \dot{\ell}(\vartheta, y) \left[ d\hat{F}_n(y) - f(\vartheta, y) dy \right], \quad x \in \mathcal{R}$$

and the corresponding statistic

$$\begin{aligned}
\hat{U}_n(x) &= \frac{\sqrt{n}}{\sqrt{I(\hat{\vartheta}_n)}} \int_{-\infty}^x \dot{\ell}(\hat{\vartheta}_n, y) \left[ d\hat{F}_n(y) - f(\hat{\vartheta}_n, y) dy \right] \\
&= \frac{1}{\sqrt{I(\hat{\vartheta}_n)}} \int_{-\infty}^x \dot{\ell}(\hat{\vartheta}_n, y) d\sqrt{n} \left[ \hat{F}_n(y) - F(\vartheta_0, y) \right] \\
&\quad + \frac{\sqrt{n}}{\sqrt{I(\hat{\vartheta}_n)}} \int_{-\infty}^x \dot{\ell}(\hat{\vartheta}_n, y) \left[ f(\vartheta_0, y) - f(\hat{\vartheta}_n, y) \right] dy \\
&= \frac{1}{\sqrt{I(\vartheta_0)}} \int_{-\infty}^x \dot{\ell}(\vartheta_0, y) dB_n(y) \\
&\quad - \frac{\sqrt{n}(\hat{\vartheta}_n - \vartheta_0)}{\sqrt{I(\vartheta_0)}} \int_{-\infty}^x \dot{\ell}(\vartheta_0, y) \dot{f}(\vartheta_0, y) dy + o(1) \\
&\Rightarrow \frac{1}{\sqrt{I(\vartheta_0)}} \int_{-\infty}^x \dot{\ell}(\vartheta_0, y) dB(F(\vartheta_0, y)) \\
&\quad - \frac{1}{\sqrt{I(\vartheta_0)}} \int_{-\infty}^{\infty} \dot{\ell}(\vartheta_0, y) dB(F(\vartheta_0, y)) \int_{-\infty}^x \frac{\dot{\ell}(\vartheta_0, y) \dot{f}(\vartheta_0, y)}{I(\vartheta_0)} dy.
\end{aligned}$$

Let us put  $F(\vartheta_0, x) = t$ ,  $F(\vartheta_0, y) = s$  and  $h(\vartheta_0, s) = \dot{\ell}(\vartheta_0, y(s))$ , where  $y(\vartheta_0, s)$  is solution  $y$  of this equation  $F(\vartheta_0, y) = s$ . Then the limit process can be written as follows

$$\begin{aligned}
U(t) &= \int_0^t h(\vartheta_0, s) dB(s) - \int_0^1 h(\vartheta_0, s) dB(s) \int_0^t h(\vartheta_0, s)^2 ds \\
&= \int_0^t h(\vartheta_0, s) dw(s) - \int_0^1 h(\vartheta_0, s) dw(s) \int_0^t h(\vartheta_0, s)^2 ds \\
&\quad - w(1) \int_0^t h(\vartheta_0, s) ds = W(\tau) - W(1)\tau - w(1) \int_0^t h(\vartheta_0, s) ds.
\end{aligned}$$

Therefore the limit statistic is not free of distribution and this approach does not allow to construct the ADF GoF test.

## References

- [1] Dachian, S. and Kutoyants, Yu.A. (2007) On the goodness-of-fit tests for some continuous time processes, in *Stat. Models for Bio.-Tech. Systems*, F.Vonta *et al.* (Eds), Boston, 395-413.

- [2] Darling, D. A. (1955) The Cramér-Smirnov test in the parametric case. *Ann. Math. Statist.*, 26, 1-20.
- [3] Durrett, R. (1996) *Stochastic Calculus. A practical introduction*. Boca Raton: CRC Press.
- [4] Fournie, E. (1992) Un test de type Kolmogorov-Smirnov pour processus de diffusions ergodique. *Rapport de Recherche*, **1696**, INRIA, Sophia-Antipolis.
- [5] Freidlin, M. I. and Wentzell, A. D. (1998) *Random Perturbations of Dynamical Systems*. 2nd Ed., Springer, N.Y.
- [6] Hwang, S.Y. and Basawa, I.V. (1993). Asymptotic optimal inference for a class of nonlinear time series models. *Stochastic Process Appl.* 46, 91-113
- [7] Iacus S., Kutoyants Yu. A. (2001) Semiparametric hypotheses testing for dynamical systems with small noise. *Math. Methods Statist.* 10, 1, 105-120.
- [8] Khmaladze, E. (1981) Martingale approach in the theory of goodness-of-fit tests. *Theory Probab. Appl.* , 26, 240-257.
- [9] Kleptsyna, M., Kutoyants Yu. A. (2013) On asymptotically distribution free tests with parametric hypothesis for ergodic diffusion processes. To appear in *Statist. Inference Stoch. Processes*, (arXiv:1305.3382).
- [10] Kutoyants, Yu.A. (1994) *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.
- [11] Kutoyants, Yu.A. (1998) *Statistical Inference for Spatial Poisson Processes*, Springer, N.Y.
- [12] Kutoyants, Yu.A. (2004) *Statistical Inference for Ergodic Diffusion Processes*, Springer, London.
- [13] Kutoyants, Yu. A., (2011) On goodness-of-fit tests for perturbed dynamical systems. *J. Statist. Plann. Inference*, 141, 1655-1666.
- [14] Kutoyants, Yu. A., (2013) On asymptotic distribution of parameter free tests for ergodic diffusion processes. To appear in *Statist. Inference Stoch. Processes*, (arXiv:1302.1026).
- [15] Kutoyants, Yu. A., (2013) On ADF GoF tests for perturbed dynamical systems. Submitted.

- [16] Kutoyants, Yu. A., (2014) On ADF goodness-of-fit tests for stochastic processes. To appear in *New Perspectives on Stochastic Modeling and Data Analysis*, J. Bozeman, V. Girardin and C. H. Skiadas (Eds).
- [17] Kutoyants Y.A. and Zhou, L.(2013) On approximation of the backward stochastic differential equation. To appear in *J. Statist. Plann. Inference*, (arXiv:1305.3728).
- [18] Liptser, R. and Shiriyayev, A.N. *Statistics of Random Processes*. v. 2, 2-nd ed. Springer, N.Y., 2005.
- [19] Maglapheridze, N., Tsigroshvili, Z. P. and van Pul, M. (1998) Goodness-of-fit tests tests for parametric hypotheses on the distribution of point processes, *Math. Methods. Statist.* 7, 60-77.
- [20] Negri, I. and Nishiyama, Y. (2009) Goodness of fit test for ergodic diffusion processes. *Ann. Inst. Statist. Math.*, 61, 919-928.
- [21] Negri, I., and Zhou, L. (2012) On goodness-of-fit testing for ergodic diffusion process with shift parameter. To appear in *Statist. Inference Stoch. Processes*, (arXiv:1203.6547).
- [22] Yoshida, N. (1996) Asymptotic expansions for perturbed systems on Wiener space: maximum likelihood estimators, *J. Multivariate Analysis*, 57, 1-36.